

## **PROJECT RAG — PROGRESS 4**

Workflow 3 (Google Drive → Embedding → Pinecone)

Nama: Janu Farras Saguna

NIM: 24523135

Tanggal: 2025-12-08

### **1. Tujuan**

Membangun Knowledge Base (Database Pengetahuan) untuk AI. Sistem ini membaca dokumen PDF dari Google Drive, kemudian memecah isi dokumen menjadi bagian-bagian kecil (chunking). Setelah itu teks diubah menjadi vector embedding menggunakan OpenAI dan disimpan ke dalam database vektor Pinecone.

### **2. Skema Data dan Arsitektur**

Sumber data berasal dari folder Google Drive yang berisi file PDF atau Markdown.

Proses dilakukan menggunakan n8n dengan tahapan sebagai berikut:

1. Download, yaitu mengunduh file dalam bentuk binary.
2. Loader, untuk membaca konten teks dari dokumen PDF atau Markdown.
3. Splitter, untuk memotong teks menjadi beberapa bagian kecil dengan ukuran sekitar 500 karakter.
4. Embedding, menggunakan model text-embedding-3-small dari OpenAI.

Data hasil embedding disimpan ke Pinecone Index dengan detail berikut:

Nama Index: honda-ai-cs

Dimension: 1536

Metric: Cosine

### **3. Artefak**

Link GitHub: <https://github.com/jaenu-dev/AI-Customer-Service-Honda>

Workflow File: /n8n/RAG Honda AI CS.json

### **4. Screenshot Bukti**

1. Tampilan dashboard Pinecone yang menunjukkan index berhasil dibuat.
2. Tampilan jumlah vektor yang masuk ke Pinecone (jumlah vektor tidak bernilai 0).
3. Tampilan eksekusi workflow n8n yang berjalan sukses (berwarna hijau).

### **5. Kesimpulan**

Database vektor berhasil dibangun. AI kini memiliki sumber pengetahuan eksternal yang berisi informasi spesifik tentang produk Honda dan siap digunakan untuk proses pencarian (retrieval) pada tahap selanjutnya.